Uncertainty and Graphical Analysis

Introduction

Two measures of the quality of an experimental result are its accuracy and its precision. An accurate result is consistent with some ideal, "true" value, perhaps a commonly accepted value from the scientific literature. When a literature value is not available, we often perform an additional measurement by other methods. Different methods are usually prone to different errors. We can hope that, if two or three different methods yield consistent results, our errors are small. However, measurements made by different methods never agree exactly. If the discrepancy is small enough, we claim that the results are consistent and accurate. Most of our work with uncertainties will address the question, "How small is small enough?"

Precision refers to the reproducibility of a result made using a particular experimental method. When random variations are large, the precision is low, and vise versa. While we should work hard to reduce the size of random effects, they cannot be entirely eliminated. When we claim that two measurements are consistent, we are claiming that their difference (the discrepancy) is smaller than these random variations. Since many quantities of interest are calculated from measured values, we also need to know how random variations in measured quantities affect the results of these calculations.

Measurements in the presence of random deviations

Mean and standard deviation of the mean[1]

In the presence of random variations, the best estimate of a physical quantity is generally given by the average, or mean. The average value of a set of N measurements of x, $(x_1, x_2, x_3, ...x_N)$, is given by

$$x_{avg} = \frac{x_1 + x_2 + x_3 + ... + x_N}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{13.1}$$

[1] A good reference for much of the information in this section is John R. Taylor, *An Introduction to Error Analysis—The Study of Uncertainties in Physical Measurements, 2nd Edition* (University Science Books, Herndon, Virginia, USA, 1987), especially Chapter 5.

The individual measurements of x will generally deviate from x_{avg} due to random errors. The standard deviation of x, denoted $\sigma(x)$, indicates how far a typical measurement deviates from the mean. The value of $\sigma(x)$ reflects the size of random errors.

$$\sigma(x) = \sqrt{\frac{(x_1 - x_{avg})^2 + (x_2 - x_{avg})^2 + (x_3 - x_{avg})^2 + (x_4 - x_{avg})^2 + \ldots (x_N - x_{avg})^2}{N-1}}$$
$$= \frac{1}{\sqrt{(N-1)}} \left[\sum_{i=1}^{N} (x_1 - x_{avg})^2 \right]^{1/2} \tag{13.2}$$

A small standard deviation indicates that the measurements (x-values) are clustered closely around the average value, while a large standard deviation indicates that the measurements scatter widely relative to the average value. Thus a small standard deviation indicates that this particular quantity is very reproducible—that is, the measurement is very precise. Note that the units of the standard deviation are the same as the units of the individual measurements, x_i.

The relation between the standard deviation to the deviation of the data from its average value is illustrated in Figure 13.1. Figure 13.1 is a histogram of 100 scores, chosen from a set of over 1000 random scores with an average was 85 and a standard deviation of 7.5. Because of their random distribution, the average of the 100 scores is not exactly 85, and their standard deviation is not exactly 7.5. Because we cannot take an infinite number of measurements, Equations 13.1 and 13.2 are only approximations to the true average and standard deviation. On average, the approximations improve as the number of measurements, N, increases.

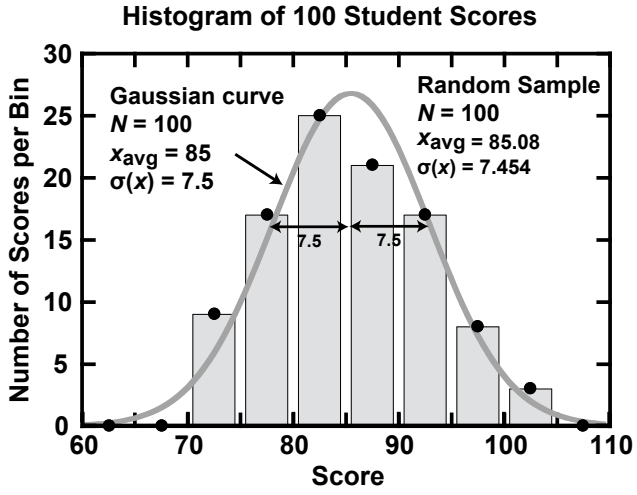

Figure 13.1. Histogram of 100 scores with an average of 85 and a standard deviation of 7.5. The smooth curve is the Gaussian function corresponding to the same number of measurements, average, and standard deviation.

The Gaussian function, $G(x)$, corresponding to 100 scores with an average of exactly 85 and a standard deviation of exactly 7.5 is also shown in Figure 13.1. According to the Central Limit Theorem

of statistics, the Gaussian function represents the ideal distribution of scores for a given N, x_{avg}, and $\sigma(x) = \sigma$ *if* the scores have a finite average and the measurements are statistically independent. These conditions apply to most of the measurements made in lab. (Important exceptions are found in the stock market, among other things.)

$$G(x) = \frac{N}{2\pi\sigma} \exp\left[\frac{-(x - x_{avg})^2}{2\sigma^2}\right] \tag{13.3}$$

The value of the standard deviation in the context of uncertainties is that the probability of finding a score at some distance from the average falls in a predictable way as the distance increases. For an ideal Gaussian distribution, 68% of the measurements lie within one standard deviation of the mean (x_{avg}). In Figure 13.1, 63 scores (63% of 100) lie within 7.5 points of 85. Ideally, 95% of the scores lie within two standard deviations (here, ± 15 points) of the average. Ideally, one would expect 99.7% of the points to lie within three standard deviations (here, ± 22.5 points) of the average. No score in Figure 13.1, is more than three standard deviations from the average. (All of the scores lie between $x_{avg} - 3\sigma = 62.5$ and $x_{avg} + 3\sigma = 107.5$.) Unless the total number of scores is very high, the probability of finding a score more than 3σ from the average is quite low.

Since the standard deviation characterizes random errors, we can pretty much rule out random errors as the source of any difference greater than 3σ. We will make this assumption in the physics labs, although the precise probabilities will usually differ from those given by the ideal Gaussian function. For instance, when the number of measurements is small, our estimates of x_{avg} and $\sigma(x)$ may be poor. In more advanced work, it can be important to correct for this lower precision.[2] When one is attempting to show that one measurement out of a large number differs significantly from the others, a higher threshold for significance (4σ or 5σ) may be necessary.

Since the result of an experiment is generally an average value, we need a measure of the precision of the average. This is called the "standard deviation of the mean," $\sigma(x_{avg})$. Although one can repeat the entire set of N measurements several times to compute $\sigma(x_{avg})$, statistics allows us to estimate $\sigma(x_{avg})$ using the original N measurements alone:

$$\sigma(x_{avg}) = \frac{1}{\sqrt{N(N-1)}}\left[\sum_{i=1}^{N}(x_1 - x_{avg})^2\right]^{1/2} = \frac{\sigma(x)}{\sqrt{N}} \tag{13.4}$$

The standard deviation function of most spreadsheet programs (Excel, OpenOffice), Capstone, and calculators gives $\sigma(x)$, from Equation 13.2. To calculate the standard deviation of the mean from this number, you must divide by the square root of N, the number of measurements.

On the other hand, spreadsheet Regression functions and Capstone's curve fit function provide the standard deviation of the mean, $\sigma(x_{avg})$ from Equation 13.4.

[2] Student's t-test is used to make this adjustment in more advanced work. This is described at the end of Chapter 5 in John R. Taylor, *op. cit.*, and in many statistics books.

Other methods for estimating the effect of random errors

When several measured quantities are used in a calculation, a relatively crude measurement of one quantity may contribute little to the overall uncertainty. If so, there is little point in improving the measurement. To demonstrate that the uncertainty is small, we must provide an upper bound on the uncertainty and show that the effect of this uncertainty is indeed relatively small.

Smallest division

Most measuring devises have a a smallest division that can be read. In this case, one can use the size of the smallest division as an upper bound on the uncertainty. In some cases, it is appropriate to to use one-half of this smallest division. For instance the smallest division displayed on a meter stick is usually 1 mm. The distance d is read to the nearest mark. Suppose, for example, you look at the meter stick a few times and read d = 85 mm each time. Because you never measured 84 or 86 mm, you are confident that $84.5 \leq d \leq 85.5$. That is, the magnitude of the uncertainty in d is less than 0.5 mm. This is a useful upper bound. You must use your judgement in cases where the measurement cannot be practically made with this precision. For instance, your precision can be much worse if you don't have a clear view of the ruler.

Interpolation

If the uncertainty in such a measurement is not small relative to the other uncertainties in an experiment, a better estimate of the uncertainty is needed. In this case, taking the standard deviation of the mean of multiple measurements is necessary. For instance, you can estimate d to one-tenth of a mm using a meter stick. (Estimating values between the marks is called interpolation.) In this case, repeated estimates, made with care, will disagree, and you can calculate the standard deviation of their mean.

Manufacturer's specification

The user manuals for many instruments (electronic ones in particular) often include the manufacturer's specification as to the "guaranteed" reliability of the readings. For example, the last digit on the right of digital voltmeters and ammeters is notoriously inaccurate. In this case, it makes sense to use the manufacturer's specification as a simple upper bound.

Terminology—Uncertainty and significant digits

Because the standard deviation is not the only measure of random variation, it helps to have another name and symbol for this quantity. We will call the the expected effect of random variation on x_{avg} its uncertainty, and represented it by the symbol $u(x_{avg})$. If the average and standard deviation of x are available, the best estimate of x is x_{avg}, and the best estimate of the uncertainty of x_{avg} is the standard deviation of its mean, $\sigma(x_{avg})$. Then $u(x_{avg}) = \sigma(x_{avg})$. The uncertainty is often indicated by a $\pm$ sign after the average value. For instance, you might specify a length measurement as "1.05 $\pm$ 0.02 mm. Because there is more than one way to estimate the uncertainty, you must also specify how your estimate was made. For instance, the result of a length measurement may be reported as "1.05 $\pm$ 0.02 mm, where the uncertainty is the standard deviation of the mean of five length

readings;" or "24 $\pm$ 1 mm, where the uncertainty is the distance between marks on the meter stick."

With or without a formal uncertainty estimate, you are expected to have a general idea of the uncertainties of the numbers you use. These uncertainties are communicated by the number of significant digits you provide with the number. For instance, a length written as 3.14 mm has an implied uncertainty of less than 0.1 mm; the inclusion of a digit in the second decimal place means that you have some knowledge of it. In your lab notebook and reports, you should not use more significant digits than are justified by your knowledge. Since rounding operations slightly increase the uncertainty in the last decimal place, it is appropriate to keep one extra significant digit in each step of a calculation. However, the final result must be rounded to an appropriate number of significant digits. Most physics texts include a discussion of significant figures.

Uncertainties in calculated quantities—the Derivative Method[3]

Derivatives can be used to estimate the uncertainty associated with a function of the measured quantity, $f(x)$, due to uncertainty in the measured variable, x. We normally have an experimental value of x_{avg}. To see how the uncertainty in x affects $f(x_{avg})$, we can plot $f(x)$ as shown in Figure 13.2. The change due to small variation in x is given by $\Delta f \approx f'(x)\Delta x$, where $f'(x)$ is the slope (and the derivative) of $f(x)$ at x_{avg}.

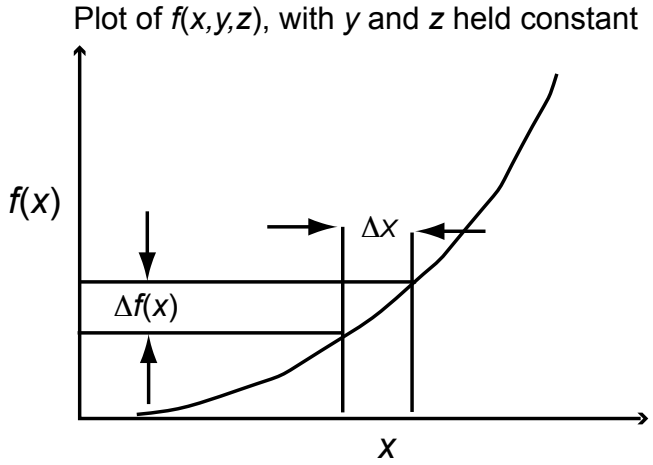

Figure 13.2. Diagram relating the uncertainty in $y = f(x)$ due to the uncertainty in x.

For the simple function $f(x) = 1/x$, with, $x_{avg} = 2.0$ and $u(x_{avg}) = 0.1$, the uncertainty in $f(x)$, $u[f(x)]$, is

$$u[f(x)] = \sqrt{\left[\frac{df}{dx}u(x_{avg})\right]^2} = \sqrt{\left[-\frac{1}{x^2}u(x_{avg})\right]^2} = \sqrt{\left(\frac{1}{4.0}\right)^2(0.1)^2} = 0.25 \qquad (13.5)$$

[3]Good references on the propagation of uncertainties include John R. Taylor, *op. cit.*, Section 3.11, "A general formula for error propagation," and Douglas C. Montgomery, George C. Runger, and Norma Faris Hubele, Section 3.12.3, "Nonlinear functions of independent random variables," *Engineering Statistics*, Fifth Edition (John Wiley, Hoboken, New Jersey, 2012).

If f is a function of more than one variable, say (x,y,z), where x, y, and z represent three measured quantities, the uncertainty in $f(x,y,z)$ is found by computing uncertainties for each variable alone and adding them in quadrature, as explained below. The uncertainty due to x is computed by treating $f(x,y,z)$ as a function of x only. Then from Equation 13.5

$$u[f(x)] = \sqrt{\left[\frac{\partial f}{\partial x}u(x_{avg})\right]^2} \tag{13.6}$$

where we introduce the ∂ symbol to indicate that the y and z variables are being treated as constants when the derivative is taken. This is equivalent to assuming that the variables are *independent*; that is, none of the variables are completely determined by any subset of the others. Likewise:

$$u[f(y)] = \sqrt{\left[\frac{\partial f}{\partial y}u(y_{avg})\right]^2} \qquad u[f(z)] = \sqrt{\left[\frac{\partial f}{\partial z}u(z_{avg})\right]^2} \tag{13.7}$$

where $u[f(x,y,z)]$ is the estimated uncertainty in $f(x,y,z)$; $u(x_{avg})$, $u(y_{avg})$, and $u(z_{avg})$ are the uncertainties in the measured values of x_{avg}, y_{avg}, and z_{avg}, respectively, all evaluated at $(x_{avg}, y_{avg}, z_{avg})$. Again, the ∂ symbols indicate that x and z are treated as constants when the derivative with respect to y is taken; likewise x and y are treated as constants when the derivative with respect to z is taken.

If you draw a two dimensional version of Figure 13.2, the Pythagorean theorem can be used to show that the uncertainties add like the edges of a right triangle, that is, in "quadrature." (This is how individual deviations add when a standard deviation is calculated.) For a function of three variables, the uncertainties add in the same way:

$$u[f(x,y,z)] = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 u(x_{avg})^2 + \left(\frac{\partial f}{\partial y}\right)^2 u(y_{avg})^2 + \left(\frac{\partial f}{\partial z}\right)^2 u(z_{avg})^2} \tag{13.8}$$

This technique can be generalized to account for as many measured parameters as necessary. When uncertainties from difference sources are added in this way, the result is called the "combined standard uncertainty,"[4] or the "standard uncertainty."[5]

Consider the function $f(x,y,z) = x^{1/2}y^2\sin(z)$. To illustrate the difference between the derivatives used to calculate uncertainties, considerthe regular (or total) derivative of $f(x,y,z)$ with respect to x, calculated using the product rule for derivatives.

$$\frac{df}{dx} = \frac{y^2\sin(z)}{2x^{1/2}} + x^{1/2}2y\sin(z)\frac{dy}{dx} + x^{1/2}y^2\cos(z)\frac{dz}{dx} \tag{13.9}$$

[4]BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML, "GUM: Guide to the Expression of Uncertainty in Measurement," Organization for Standardization (2008).

[5]Barry N. Taylor and Chris E. Kuyatt, "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results," NIST Technical Note 1297, 1996 edition (National Institute of Standards and Technology, Gaithersburg, Maryland, 1994).

However, in an experiment, x, y, and z are independent variables. Therefore we expect $dy/dx = dz/dx = 0$. For the purposes of calculating the contribution of $u(x_{avg})$ to the uncertainty of f, y and z might as well be constants. The three required derivatives of $f(x,y,z)$ from Equation 13.8 are:

$$\frac{\partial f}{\partial x} = \frac{y^2 \sin(z)}{2x^{1/2}} \qquad \frac{\partial f}{\partial y} = 2x^{1/2} y \sin(z) \qquad \frac{\partial f}{\partial z} = x^{1/2} y^2 \cos(z) \tag{13.10}$$

In practice, taking derivatives can be a lot of work. However, many calculations involve products, which are simplified by starting with the natural logarithm of the calculated quantity. Since

$$\frac{\partial}{\partial x}[\ln(f)] = \frac{1}{f}\frac{\partial f}{\partial x} \quad , \tag{13.11}$$

we can calculate the derivatives we need from the derivatives of the logarithm. Since the logarithm function splits our function into terms with simple (partial) derivatives, they are easy to compute. In our example, $\ln(f) = (1/2)\ln(x) + 2\ln(y) + \ln(\sin z)$, so

$$\begin{aligned} \frac{\partial}{\partial x}[\ln(f)] &= \frac{\partial}{\partial x}\left[\frac{\ln(x)}{2}\right] = \frac{1}{2x} \\ \frac{\partial}{\partial y}[\ln(f)] &= \frac{\partial}{\partial y}[2\ln(y)] = \frac{2}{y} \\ \frac{\partial}{\partial z}[\ln(f)] &= \frac{\partial}{\partial y}[2\ln(\sin z)] = \frac{\cos z}{\sin z} = \cot(z) \end{aligned} \tag{13.12}$$

Substituting these partial derivatives into Equation 13.8 yields

$$\frac{u[f(x,y,z)]}{f(x,y,z)} = \sqrt{\left[\frac{u(x_{avg})}{2x_{avg}}\right]^2 + \left[\frac{2u(y_{avg})}{y_{avg}}\right]^2 + [\cot(z_{avg})u(z_{avg})]^2} \tag{13.13}$$

While this expression is not pretty, it is much simpler than the one obtained by substituting the derivatives of Equation 13.10 directly into Equation 13.8. For simplicity, the uncertainty is in Equation 13.13 is expressed as a fraction of the value of $f(x,y,z)$. This is called the "relative uncertainty," or more completely, the "relative combined standard uncertainty."

Using uncertainties to compare measurements or calculations

Suppose you have measured a cart's mass, $m_{F/a}$, from force and acceleration measurements and Newton's Second Law, $F = ma$. To check for systematic errors, you have also measured the cart's mass using an electronic balance, with the result m_{bal}.

A straightforward way to determine whether these two measurements is to compare the discrepancy between the two measurements, say $\Delta = \left|m_{F/a} - m_{bal}\right|$, with the expected uncertainty of Δ,

that is $u(\Delta)$. As illustrated in Figure 13.1, the probability of Δ being more than three standard deviations from the mean because of random errors alone is quite small. Therefore, if $\Delta > 3u(\Delta)$ most of the discrepancy is almost certainly due to systematic problems. In this case, we say that the measurements of $m_{F/a}$ and m_{bal} are not consistent.

The ratio between the discrepancy and its combined standard uncertainty is a useful measure of the seriousness of a discrepancy. Because this ratio is similar to the t-statistic of classical statistics, we call it the t'-score.[6] In this example,

$$t' = \frac{\Delta}{u(\Delta)} = \frac{\Delta}{\sqrt{u(m_{F/a})^2 + u(m_{bal})^2}} \tag{13.14}$$

When you compare experimental results and find $t' > 3$, you should carefully review your calculations and measurement procedures for errors. If systematic errors appear to be significant, and you know what they might be, you should describe them in your lab notes. If time permits, repeating a portion of the experiment is in order. Whatever your conclusion, your lab notes must indicate how you estimated your uncertainties.

In the United States, the general authority on the reporting of uncertainties is the National Institute of Standards and Technology.[7] These standards have been developed in consultation with international standards bodies. When the potential consequences of a decision are critical or when the data are unusual in some way, one should consult a statistician.[8]

Determining functional relationships from graphs

Linear relations are simple to identify visually after graphing and are easy to analyze because straight lines are described by simple mathematical functions. It is often instructive to plot quantities with unknown relationships on a graph to determine how they relate to one another. Since data points have not only measurement uncertainties but also plotting uncertainties (especially when drawn by hand), slopes and such should not be determined by using individual data points but by using a "best-fit line" that appears to fit the data most closely as determined visually. If graphing software is used, then the slope of the line can usually be determined by a computer using a "least squares" technique. We won't go into detail about these methods here.

Linear functions $(y = mx + b)$

If x and y are related by a simple linear function such as $y = mx + b$ (where m and b are constants), then a graph of y (on the vertical axis) versus x (on the horizontal axis) will be a straight line whose slope ("rise" over "run") is equal to m and whose y-axis intercept is b. Both m and b can be

[6] N. T. Holmes and D. A. Bonn, "Quantitative comparisons to promote inquiry in the introductory physics lab," Phys. Teach. **53**(7), 352 (2015). DOI: 10.1119/1.4928350

[7] Ibid, Barry N. Taylor and Chris E. Kuyatt.

[8] W. Edwards Deming, *Out of the Crisis* (MIT Press, Cambridge, Massachusetts, 1982). Some authors attribute the ability of Japanese automakers to break into the U.S. market to their skillful application of the principles of statistical quality control popularized by W. Edwards Deming and Joseph Juran.

determined once the graph is made and the "best-fit" line through the data is drawn. if $x = 0$ does not appear on your graph, b can be found by determining m and finding a point (x,y) lying on the "best-fit" line; then equation $y = mx + b$ can be solved for b.

Simple power functions $(y = ax^n)$

In nature we often find that quantities are related by simple power functions with $n = \pm 0.5, \pm 1, \pm 1.5, \pm 2$, etc., where a is a constant. Except for $n = +1$, making a simple graph of y (vertical axis) and x (horizontal axis) for simple power functions will yield a curved line rather than a straight line. From the curve it is difficult to determine what the actual functional dependence is. Fortunately it is possible to plot simple power functions in such a way that they become linear.

Starting with the equation $y = ax^n$, we take the natural logarithm of each side to show

$$\ln(y) = \ln(ax^n) = \ln(a) + \ln(x^n) = \ln(a) + n\ln(x) \tag{13.15}$$

If $\ln(y)$ is plotted on the vertical axis of a graph with $\ln(x)$ plotted on the horizontal axis (This is often called a doubly logarithmic, or log-log graph.), then Equation 13.15 leads us to expect that the result is a straight line with a slope equal to n and a vertical axis intercept equal to $\ln(a)$. If the relationship between y and x is a simple power law function, then a graph of $\ln(y)$ as a function of $\ln(x)$ will be linear, where the slope is n, the power of x, and the intercept is the natural logarithm of the coefficient a. This is quite useful, because it is easy to determine whether a graph is linear. If we suspect a simple power function relationship between two quantities, we can make a log-log graph. If the graph turns out to be linear, then we are correct in thinking that it should be a simple power function and can characterize the relationship by finding values for n and a.

Exponential functions $(y = ae^{bx})$

Radioactive decay, the temperature of a hot object as it cools, and chemical reaction rates are often exponential in character. However, plotting a simple graph of y (on the vertical axis) and x (on the horizontal axis) does not generate a straight line and therefore will not be readily recognizable. A simple graphical method remedies this problem. Starting with an equation for the exponential function, $(y = ae^{bx})$. We can take the natural logarithm of each side to show

$$\ln(y) = \ln(ae^{bx}) = \ln(a) + \ln(e^{bx}) = \ln(a) + bx \tag{13.16}$$

If $\ln(y)$ is plotted on the vertical axis and x is plotted on the horizontal axis (This is called a semi-log graph.), Equation 13.16 takes the form of a straight line with a slope equal to b and a vertical axis intercept equal to $\ln(a)$. Thus any relationship between two variables of this simple exponential form will appear as a straight line on a semi-log graph. We can test functions to check whether they are exponential by making a semi-log graph and seeing whether it is a straight line when plotted this way. If so, the values of a and b that characterize the relationship can be found.

Using error bars to indicate uncertainties on a graph

When plotting points (x,y) with known uncertainties on a graph, we plot the average, or mean, value of each point and indicate its uncertainty by means of "error bars." If for example the uncertainty is primarily in the y quantity, we indicate the upper limit of expected values by drawing a bar at a position y_{max} above y_{avg}, that is, at position $y_{max} = y_{avg} + u(y_{avg})$. Similarly, we indicate the lower limit of expected values by drawing a bar at position $y_{min} = y_{avg} - u(y_{avg})$. Figure 13.3 shows how the upper error bar at y_{max} and the lower error bar at y_{min} are plotted. If the quantity x also has significant uncertainty, one adds horizontal error bars (a vertical error bar rotated 90°) with the rightmost error bar at position x_{max} and the leftmost error bar at position x_{min}.

Occasionally one encounters systems where the upper and lower error bars have different lengths. In this case, the upper uncertainty, $u_+(y_{avg})$ does not equal the lower uncertainty, $u_-(y_{avg})$.

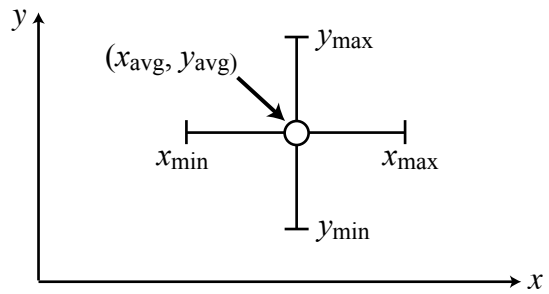

Figure 13.3. Diagram of error bars showing uncertainties in the value of the x- and y-coordinates for point (x_{avg}, y_{avg}) When you print a graph in lab, the labels are omitted.